Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

ON THE STUDY OF STATISTICS.[1]

I.

THE word statistics in English has a narrow and well-defined meaning. It means the enumeration or counting of something. By the statistics of imports and exports we understand the enumeration of all the commodities imported and exported, with their values. Bank statistics are the figures showing the transactions of the banking institutions of a country, the loans and discounts, the circulating notes and the specie reserve — in short, an exposition in numbers of the business of banks. A statistical abstract is a collection of statistics of various kinds pertaining to a single country or group of countries.[2] A statistical manual or handbook is a collection in condensed form of such figures for a country or group of countries; and a dictionary of statistics is the same material arranged alphabetically. So common and popular is this use of the term, that speaking of statistics involuntarily brings to mind volumes filled with statistical tables, pages of figures — raw material of such crudeness that almost any use can be made of it by skillful manipulation. Hence also the common sneer that statistics can be made to prove anything, that there are statistics on both sides or on every side of a question.

The cultivators of statistics have themselves set the seal of their approbation on this narrow interpretation of the word. Professor Ingram, when president of Section F (Sociology and Statistics) of the British Association for the Advancement of Science, said in his opening address:

> It is plain that though statistics may be combined with sociology in the title of Section F, the two cannot occupy a coördinate position.

---

[1] Statistik und Gesellschaftslehre. Von Dr. Georg von Mayr. Erster Band: Theoretische Statistik. Aus Handbuch des Oeffentlichen Rechts. Leipzig, J. C. B. Mohr, 1895. — large 8vo., 202 pp.

[2] *E.g.*, the Statistical Abstract of Great Britain and Ireland; of The Colonies; of Foreign Countries.

> For it is impossible to vindicate for statistics the character of a science; they constitute only one of the aids or *adminicula* of science.

And Mr. Wynnard Hooper, in a paper read before the London Statistical Society and later in the article "Statistics" in the *Encyclopaedia Britannica,* said that by statistics we always mean the statistics of something, as of the number of bags of coffee imported into London during a certain week, or the number of passengers carried by the Great Northern Railway in a year. The members of the society present when the paper was read were practically unanimous in agreeing that Mr. Hooper had proved his point.

It must be admitted that the official statisticians in England and America very largely accept this view. Their work consists (according to their own opinion) in gathering statistics of something, that is, in mere enumeration and in publishing the result of the enumeration, often without note, explanation or comment. They regard themselves as blind instruments, transmitters of intelligence, like the telephone that receives and carries the message without understanding the content. They occupy a position similar to that of the Scriptural writers under that theory of inspiration according to which they simply wrote down what they were told to write, the message being true not because they had had anything to do with it, but precisely because they had had nothing to do with it.

It is of course obvious that the official statistician should be perfectly honest and impartial in his work. He must give us the statistics as he finds them, not as he thinks they ought to be or desires that they might be. His primary object is not to prove a thesis but to furnish material by which theses may be proved. His attitude is to be objective, and the statistics are to reflect facts, not his subjective opinions. It is perhaps the recognition of this that has led statisticians to leave their statistics in a somewhat crude and undigested form. The attitude is certainly praiseworthy; and although this professed modesty may sometimes have been used to cover indolence or lack of technical skill, yet it has been a step in the direction of scientific progress.

There are signs, however, that this perfect passivity on the part of the statistician is not all that is demanded of him. He must present his material in an impartial way; but that does not preclude its being presented in a systematic way. He must utilize his experience to make his figures intelligible. He is the one who best knows how to manipulate them and it is his duty to make them manageable. His familiarity with the original material gives him expert knowledge which saves him from mistakes and misinterpretations such as an outsider may easily commit. It is the business of the statistician to place this knowledge at the service of the public and to give them not merely statistics, but intelligible statistics.

Advances in the direction indicated are seen in the care with which in modern statistical publications not only the absolute but the relative figures are worked out for comparison. If it is the statistics of sex, we have the proportionate number of males and females in different countries, at different times, the relative increase or decrease during the intercensal period, the proportion at different age-periods, among different classes of the population, *etc.* A further evidence of progress is the use of diagrams, cartograms and other methods of appealing to the eye. All this is merely technical. It is the work of the mechanician in preparing his instrument for its work. It requires skill, but it is merely technical skill. It is still statistics of "something," but it is statistics in a refined form.

The modern statistician does not stop here. He elucidates what his statistics show and what they do not show. He gives us the advantage of his expert knowledge so that we shall not misinterpret the figures. An example or two will make this clear. The bulletin of the eleventh census giving the number of insane in the United States reported a larger number of inmates of insane asylums in 1890 than in 1880; the superintendent of the census remarked that this was due to the larger number of such institutions and the increased care for such unfortunates. The statistician of the Treasury Department recently showed an error of $75,000,000 in the returns of values of imports of a previous year, due to the fact that imports from countries with a depreciated paper currency had not been valued on a specie basis as they should have been. Occupation statistics always require explanation because of the difficulties of classification. A dairy maid — is she an agricultural laborer or a domestic? A clergyman who teaches school — is he a clergyman or a teacher?[1]

Very often the statistician finds himself compelled to go further in his elucidation. He tries to explain his figures. The English census finds more old women than old men, especially in cities. The general fact, the editor thinks, is due to the less dangerous character of the occupations carried on by women, and the secondary fact, to the circumstance that in cities there is greater opportunity open to old women than to old men in the way of light employments. The number of blind in Ireland is decreasing, owing to the better medical service in the poor-houses and asylums. The distribution of the foreign-born in the United States has been influenced largely by the mere historical fact of the prevalence of slavery in the South. All these things the statistician feels himself bound to point out lest the figures be misinterpreted.

Finally, the statistician often proposes to himself certain problems, and tells us whether his figures can give the solution or not. For instance, the problem of over-crowded tenement houses. Can we get a statistical measurement of the evil so as to determine its prevalence and compare different countries, cities or districts? On this point, the editor of the English census of 1891 speaks as follows:

> We must first define what we are to take as a standard of over-crowding. It is plain that the number of rooms and occupants is not in itself an absolutely sure guide, because rooms differ largely from each other in size. Still we may be tolerably certain that the rooms in tenements with less than five rooms will not in any but exceptional cases be of large size, and that ordinary tenements which have more than two occupants per room, bedrooms and sitting-rooms included, may safely be considered as unduly over-crowded. [On this basis] there were 481,653 over-crowded tene-

---

[1] See Census, England, Part IV, p. 35, for detailed discussion.

ments in which dwelt 3,258,044 persons, or 11.23 per cent of the total population; the average number of persons per room amounting to 2.8.[1]

All this is evidently more than the "statistics of something." The statistician is more than the enumerator of facts. He is a student of the relations and regularities of social life. He is more than the mechanician. He is an expert, manipulating a certain instrument of investigation for the purpose of accomplishing certain results. He is not a sociologist, in the full sense of the word, because he does not seek the whole explanation of social organization, but deals only with a certain range of facts.

Various uses of the word statistics show that this more extended activity of the statistician is justifiable and expected. When we say that "statistics prove," we have in mind not only an array of figures, but comparison of these figures with others in order to show that certain relations exist. When we say that "statisticians assure us" that an increase of food-prices is followed by an increase of crimes against property, we refer to an elaborate study of the changes in these two phenomena relatively to each other, extended over many countries and years. We sometimes use the term statistical method, and occasionally even statistical science. On the other hand, we do not use the term statistics in the singular, as we use mathematics, ethics and politics. The Germans have *die Statistik* and the French *la statistique,* denoting the method, art or science. It is evident that in English we have the thing if not the name. How shall we define the procedure and how limit the activity of the statistician? To exactly what work should the statistician address himself, and what is the relation of his work to that of the sociologist? The growing interest in the study of sociology in this country makes these questions pertinent.

On the continent of Europe the statisticians have busied themselves with these questions for many years, and we have a considerable body of literature on the theory of statistics. Beginning with Quetelet, we have Wagner, Knies, Engel, Von Mayr, Knapp and Lexis in Germany; Guerry, Block and Levasseur in France; Gabaglio, Morpurgo and Messedaglia in Italy; and many others who have written on the subject. In the important work to which I refer at the beginning of this article Von Mayr reiterates his own theory, elaborated in the light of the work of his predecessors and contemporaries. It may, therefore, be taken as an authoritative statement of the most recent notions of the continental school in regard to statistics and their place in the study of social science. It will be worth our while to analyze these conceptions, and to try to determine if they will help us in answering the questions noted above.

---

[1] Census of England, Part IV, p. 21. For equally interesting discussions of the tenement-house question, see Mass. Bureau of Statistics of Labor, 1891 and 1892, and Census of Ireland, 1891, Part II, p. 9.

II.

With what sort of a discipline, method, art or science have we to do when we speak of statistics in the singular?

Von Mayr's answer to this question is a very simple one:

Man may be looked upon as belonging to nature, or, owing to his spiritual development, as separate from or in contrast to nature. Under the first aspect he is an object of study for natural science; the scientific investigation of the phenomena of man's activity in the sphere of his spiritual life constitutes the content of the social sciences.[1]

The scientific study of man, whether as an object of natural or of social science, may be directed to the individual, either by careful investigation of his physical body or by observation of the development of mind, and either by introspection or by study of others.

Scientific study may also be concentrated on specific (*abgegrenzte*) masses of men, simply as masses or in respect to the new relations which spring up among men living together.

---

[1] The distinction is only a formal one; for in reality man, in spite of his spiritual activity, remains a product of nature, and the natural processes which bind man down to the laws of the animal world are at the same time an important limitation of, and the necessary substratum for, his spiritual activity. The distinction is a convenient abstraction for the division of labor.

For the individuals in such masses do not merely stand like isolated atoms alongside of one another, but develop fixed relations, and especially groups or associations. These associations vary from the simplest almost unorganized groups for general social purposes to the highly organized associations for specific purposes, the most highly organized of which is the state. The object of social science is to study these masses, to observe their activity and the effect of such activity as it discloses itself in fixed relations, associations and institutions. What method shall we employ for this purpose?

Natural science can in many instances not only isolate its object but even subject it to experiment. To social science this experimental method is as a rule impossible. It can only strive by observation to obtain a comprehension of social masses, their activity and the results of such activity. This is done most completely by an exhaustive enumeration and measurement of the elements of the social masses. Such scientific observation of social masses is what we mean by statistics. Statistics may be defined, therefore, as the science of social masses.

The science of statistics as thus defined has two important sides, *viz.*, method and content, or theoretical and practical statistics. The first embraces method in its widest extent, including administrative statistics and the history of statistics. The statistician or scientific investigator must begin with the choice of material. It is impossible to observe all the facts of social life. Some it is impracticable to observe, as the facts of domestic life; others are not worth the trouble; still others it is not wise to question people about, as religious convictions. The amount of material which can be collected is limited by the resources or disposition of the administration. Scientific workers and practical statisticians should unite in determining what sort of material should be chosen.

The real questions of technique begin with the actual work of collecting the statistics. Here there are many questions requiring expert knowledge, such as the arrangement of the schedules, the periods of time chosen for the observations, the inquiries to be made, the organization of the office, the employment of other officers of the administration, the use of other material, *etc.*, *etc.* The statistician must be on his guard against various hindrances to correct returns, such as a disposition to exaggerate (centenarians); or to underestimate (tax returns); or to evade the question (physical or mental infirmities); or to answer untruthfully (previous criminality). The statistician must learn to arrange his questions in such a way that the answer to one will control the answer to another. A constant exercise of the critical faculty is necessary in order to prevent deception and imposition. The statistician must be an expert in such matters.

The next step, tabulating the figures, requires equal skill and expert knowledge. The arrangement of the figures, the choice of classes, the use of proportional rates, averages, weighted averages, median values, index numbers, diagrams, cartograms and graphical representations, — all these expedients for making the figures intelligible belong to the technique of statistics.

Finally, the statistician must arrange his tables so that they may be of scientific use for purposes of comparison and analysis. Thus, he must give the birth-rates for successive years, the density of population for various divisions of space, *etc.* It is impossible to give even a résumé of the thorough and exhaustive analysis of statistical method given in Von Mayr's book, which discloses the practical statistician as well as the trained scientist. All this pertains to statistics as a scientific method.

The practical part of the science of statistics comprises the actual knowledge thus far gained of the social masses. Von Mayr's second volume is to be devoted to an outline of such results. It will include population statistics, criminal, educational, economic and political statistics. It will give a complete picture of the social masses so far as we are able to observe them by the statistical method. Method and content together constitute the science of statistics according to the continental notion.

III.

If we acquiesce in this notion of statistics as the science of social masses, with method and content as noted above, two difficulties at once present themselves: (1) What is the relation of statistics to the other social sciences? and (2) How shall statistics divide the field with sociology? Von Mayr has an answer to these questions also, and it is as follows:

The whole field of the social sciences may be divided between (1) the general social sciences, (2) the special social sciences and (3) history. The general social sciences are those which busy themselves with social masses or associations in general. Statistics and sociology come under this head. The special social sciences deal either with certain directions of social activity or with certain special manifestations of social life. An example of the first kind is political economy; of the second is philology. Statistics occupies a peculiar position towards all of these special social sciences. It first of all studies the social masses as pure process of nature, as in the statistics of sex, births, deaths and physical characteristics. In these respects it has many relations with natural science, especially anthropology. In the second place, statistics furnishes material for all the special social sciences, especially political economy. In political science it is also extremely helpful, inasmuch as administrative statistics are often used to guide state action. Statistics is, then, the fundamental and general social science, observing social masses directly and furnishing material to the other sciences.

The second question is a more difficult one, *viz.*, as to the relation between statistics and sociology. The latter is often looked upon as the general social science, and thus seems to cover the same field as statistics. Sociology deals directly with social masses or with societies from the primitive horde to the most highly developed nations of modern times. Statistics seems to cover the same field, while not using all the material which sociology claims to utilize. Von Mayr admits that social masses (societies) can be studied as groups or associations, and that there sociology has a field of its own. He doubts, however, whether societies can be observed accurately except with the aid of statistics. Statistics deals, therefore, with the study of social masses as such; sociology studies societies according to their forms and the laws of their development, and especially with the influence of social organization on the individuals concerned. Sociology he would call the theoretical social science; statistics, the practical social science. Both are introductory to the study of the special social sciences.

It must be confessed that this distinction between statistics and sociology is not very clear. They seem to cover the same ground and deal with the same material. Von Mayr claims that part of the confusion is due to the sociologists, who are not yet agreed as to the scope of their science, nor quite clear as to their method. They claim that it is purely inductive; but what they really do (according to Von Mayr) is to find a basis in history and in the observation of society for certain subjective notions of their own.

A solution of the difficulty seems to me to be suggested by Von Mayr himself when he calls statistics the practical social science, or the exact social science (*die exacte Gesellschaftslehre*).[1] Statistics is that part of sociology which rests on exact enumerations and measurements. Statistical science consists in the exact observation of social masses and the clear presentation, by means of enumeration and comparison, of the facts and relations of social life. It is not of much consequence in this connection whether we use in English the term statistical method or statistical analysis or statistical science. It would be convenient on some acccounts if we could use the simple term statistics to include both the method and the systematized results of the method applied to human life in society. As I have already pointed out at the beginning of this essay, the thing is already there, and it would be well to attach the conception to a definite name.

The advantages of the recognition of this conception are twofold. In the first place it would legitimatize and encourage the practical statisticians in the effort to treat their material according to the best technique and to present their results in a systematic and scientific form. The men who are to perfect the method must have some intelligent comprehension of the results to be attained. The statistician (to repeat) must be something more than a mere mechanic. He must be an investigator. What we need is not so much a greater body of statistics, as better statistics. The above conception gives the statistician ample scope for the application of the most refined methods, while at the same time unifying all statistical effort by giving it a common object, *viz.*, systematic observation for the purpose of gaining knowledge of social organization.

---

[1] Von Mayr seems to confine this notion to the practical part of statistics. It seems to me impossible or at least undesirable to separate method from content.

Even more important than this would be the influence of such a conception on the student of sociology. Statistics is one of the most important instruments of investigation in that science. The sociologist must not only use statistical results, but he should be familiar with statistical methods in order to test the trustworthiness and adequateness of his material. The statistical laboratory, where he can encounter the problems of statistical observation, tabulation, analysis and induction, is as necessary for the student of sociology as the dissecting room is for the physician or field work for the geologist. It is only by such work that we can counteract the tendency to sentimentalism and *a priori* speculation, which are so common among many so-called sociologists. It is here that we can insist upon precise observation, rigid technique and provisional induction, which are the foundation of scientific method. The study of the science of statistics, both in its method or technique as the science of social masses, and in its results as the practical or exact social science, affords the best promise for the development of a sane and intelligent sociology.

RICHMOND MAYO-SMITH.